(54) Titre : PIEGE A GENES OBTENU PAR RECOMBINAISON POUR IDENTIFIER ET ISOLER DES GENES
(54) Title: GENE-TRAPPING CONSTRUCT FOR THE IDENTIFICATION AND ISOLATION OF GENES
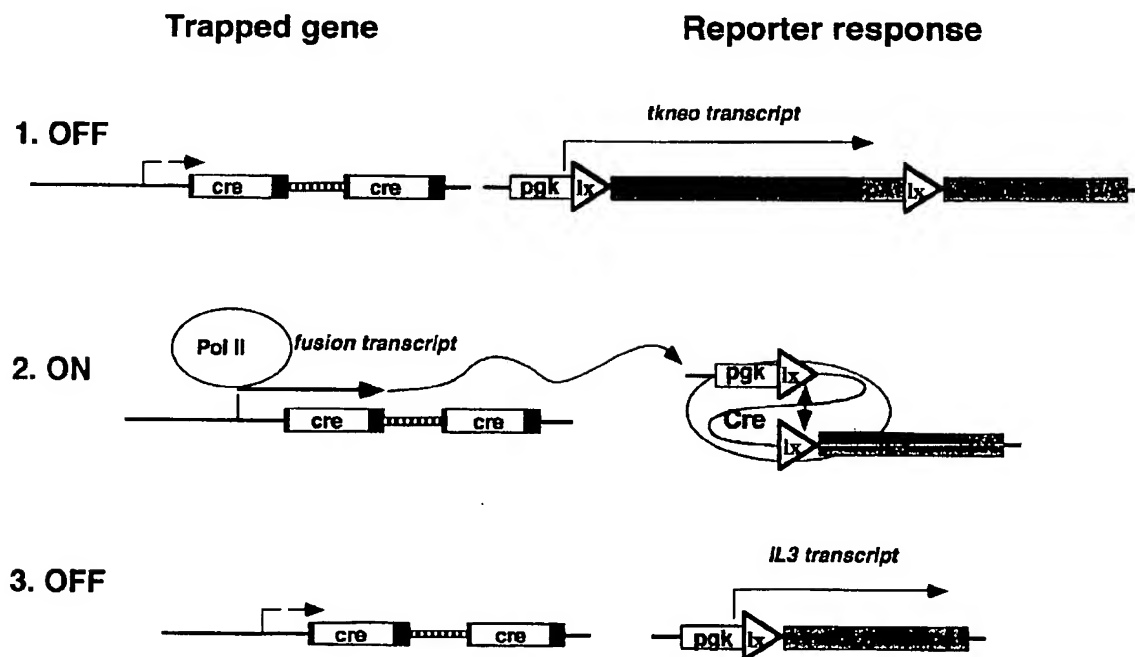
(57) Abrégé/Abstract:

The invention relates to a gene trap construct containing a first reporter gene, which, after activation, can activate a second reporter gene, and the use of this gene trap construct for identification and/or isolation of genes, in particular transiently expressed genes. The invention relates further to a cell, preferably a mammal cell, containing the above-mentioned gene trap construct. The invention relates furthermore to the use of this mammal cell for identification and/or isolation of genes, in particular transient genes. In addition, the invention relates to a vector containing the above-mentioned gene trap construction, as well as a kit for identifying and/or isolating genes, in particular transient genes, which contains at least the above-mentioned gene trap construct or the above-mentioned vector. Finally, the invention relates to a method for the identification and/or isolation of genes, in particular transient genes.

ABSTRACT

The invention relates to a gene trap construct containing a first reporter gene, which, after activation, can activate a second reporter gene, and the use of this gene trap construct for identification and/or isolation of genes, in particular transiently expressed genes. The invention relates further to a cell, preferably a mammal cell, · containing the above-mentioned gene trap construct. The invention relates furthermore to the use of this mammal cell for identification and/or isolation of genes, in particular transient genes. In addition, the invention relates to a vector containing the above-mentioned gene trap construction, as well as a kit for identifying and/or isolating genes, in particular transient genes, which contains at least the above-mentioned gene trap construct or the above-mentioned vector. Finally, the invention relates to a method for the identification and/or isolation of genes, in particular transient genes.

Gene-trapping construct for the identification and isolation of genes

The invention under consideration concerns a gene-trapping construct containing a first reporter gene which after activation can activate a second reporter gene, and the use of this gene-trapping construct for identification and/or isolation of genes, especially transiently expressed (transient) genes.

In addition the invention under consideration concerns a cell, preferably a mammalian cell, which contains the abovementioned gene-trapping construct. Besides this the invention under consideration concerns the use of this mammalian cell for identification and/or isolation of genes, particularly transient genes. Furthermore the invention concerns a vector containing the abovementioned gene- trapping construct, as well as a kit, containing at least the abovementioned gene-trapping construct or the abovementioned vector, for the identification and/or isolation of genes, particularly transient genes.

In conclusion the invention under consideration concerns a process for the identification and/or isolation of genes, particularly transient genes.

For some years one of the prime objectives of gene technology has been the isolation and identification of genes. For this there is available a whole range of procedures which have been directed in effect at the isolation and identification of permanently expressed genes.

The isolation and/or identification of genes which are expressed only transiently, as for example the genes responsible for programmed cell death, cell-cycle genes, DNA repair genes and differentiation-specific genes, is much more difficult.

To identify such genes in mammalian cells, where genetic analysis by the use of *Drosophila melanogaster* is unsuitable, a process has been developed which is based on the induction of gene fusions between a reporter gene without promotor and the control elements of a cellular gene via specific vectors, which are described as "gene traps" or "promotor traps". Various types of vector have been developed for insertion mutagenesis in mammals, whereby a reporter gene is inserted into the chromosome at a large number of places depending on chance, including in transcriptionally active areas. During selection for gene expression, clones are retained in which the

reporter gene is fused with the regulatory elements of an endogenous gene. In this way the vectors act as gene traps and provide a very helpful means for the analysis of gene function (Reviews in Hill & Wurst, 1993; Hill & Wurst 1993; von Melchner et al 1993; von Melchner & Ruley, 1995). In the majority of cases the gene-trapping vectors are transduced as recombinant retroviruses, although electroporated DNA is also used. The retroviruses display the advantage that they integrate in several areas throughout the entire genome and hence scarcely damage the neighbouring DNA (Varmus, 1988; Coffin et al., 1989; Goff, 1990; Sandmeyer et al., 1990; Withers/Ward et al., 1994).

It could be demonstrated that the gene traps are a practical means of analysing gene function in mice. Since totipotent mouse ES cells are used as cellular targets, mouse strains displaying inactivated gene function on account of mutations can be constructed. Unlike with gene splitting due to homologous recombination, the gene-trapping processes are not confined to genes, are sustainable for the cloned sequences and hence represent a process for the isolation and identification of genes as yet unknown.

Nevertheless, in order to identify and isolate genes which must first be induced in cells, that is, which are not continuously being transcribed, for example transient genes, such as genes responsible for programmed cell death, cell cycle genes, DNA repair genes and differentiation-specific genes, an additional process is necessary, in order to conclusively identify the transient cellular promotor captured in the gene trap, for instance through a durable signal independent of the promotor activity. Approximately 50% of the genes which are switched off in ES cell lines following infection/electroporation by gene-trapping vectors, manifest recessive phenotypes in mice (Friedrich and Soriano, 1991; Skarnes et al., 1992; von Melchner et al., 1992). This frequency is ten times greater than that observed following accidental insertion of retroviruses or microinjection of DNA (Gridley et al., 1987; Jaenisch et al., 1985). Due to this high efficiency of gene inactivation it appears sensible to isolate cell lines which display integration in most of the genes expressed (2-4 $\times 10^4$). It is not very practical, however, to transfer all mutations into the original line; furthermore many mutations result in genes of more limited significance. Consequently it would be desirable to exchange mutagenized ES cell clones for mutations relevant to important biological processes or genes.

With reference to this, those strategies are especially suitable which involve the preliminary selection of ES cell clones from interesting in vitro mutations, and for this employ a reporter gene to identify mutations e.g. in differentiation-specific genes. Although cultivated ES cells might have expressed genes which are not expressed in vivo, in 12 cases fusion genes were found which were expressed in ES cells as well as in embryos (De Gregory et al., 1994; Reddy et al., 1992). It further appears to be the case that a change in the expression of fusion genes during differentiation in vitro very closely predicts the change in expression during in vivo development (Reddy et al., 1992). This is in general agreement with earlier observations, in which in vitro-regulated genes were also strictly regulated in vivo (Muthuchamy et al., 1993; Rappotee et al., 1988; Rogers et al., 1991; Scholer et al., 1990; Sharpe et al., 1990).

It was therefore intended according to the invention under consideration to prepare a gene-trapping construct which made possible the isolation and identification of genes, especially transient genes, in a particularly advantageous way. A further purpose in respect to the invention under consideration was to provide a cell as well as a vector which maintained this gene-trapping construct, together with a kit for the identification and/or isolation of genes.

In conclusion, it was an intention according to the invention under consideration to present a process for the identification and/or isolation of genes, especially transient genes.

These purposes are satisfied by the present invention.

The preferred types of operation with respect to the invention under consideration are presented in the claims below.

A "reporter gene" here signifies a nucleic acid sequence which gives rise on transcription to a recognizable signal such as for example a protein product. The "first" reporter gene is regulated through cellular transcription signals in whose neighbourhood it is integrated. Once the first reporter gene is activated, the first reporter gene product leads to activation of the "second" reporter gene, in which this second activation is permanent, that is, independent of the transcriptional activity of the DNA region controlling the first reporter gene. The second reporter gene codes for a demonstrable gene product, such as for example an enzyme, demonstrable by a colour reaction, or a growth factor, which can then be selected for. The result of the combination of first and second

4

reporter genes is that even an only transiently expressed promotor can be indicated by means of a permanent signal, the second reporter gene product. The first and second reporter genes may exist in a common construct, or they may be available in various devices present in various places in the cell. The gene-trapping construct according to the invention is consequently here in principle a combination of a conventional gene-trapping construct, such as for example that described in US 5364783, corresponding to the first reporter gene construct, and a further device which can provide a permanent product independently of the promotor activity, in the present instance the second reporter gene construct.

According to one aspect of the invention under consideration, a gene-trapping construct is presented containing a first reporter gene which following activitation can activate a second reporter gene. The first reporter gene preferably codes for a recombinase, such that this recombinase is in particular preferentially Cre or Flp.

In a favoured operational form the second reporter gene may code for a protein factor, in a further favoured operational form the second reporter gene codes for IL-3 and in a particularly favoured further operational form for E. coli beta-galactosidase (lacZ).

The second reporter gene is preferably activated by deletion by the recombinase of a DNA fragment situated next to the second reporter gene, which thereby places the second reporter gene immediately downstream of a promotor permanently under its control.

The deleted DNA fragment may preferentially be an antibiotic-resistance gene, by means of which the deleted DNA fragment codes in a particularly favoured operational form for thymidine kinase-neomycin phosphotransferase fusion protein, and in a further particularly favoured operational form codes for a neomycin phosphotransferase.

The promotor before which the second reporter gene is placed is preferably the phosphoglycerate kinase promotor.

In a favoured operational form the deleted DNA fragment is flanked by sequences, which may for example be loxP or frt, targeted by the recombinase. A possibility also exists that the target sequences may be located in the U3 and/or the U5 region.

The gene-trapping construct described above can be used for identification and/or isolation of genes, particularly transient genes. Included under this particularly is the isolation and/or identification of genes responsible for programmed cell death, cell cycle genes, DNA repair genes and differentiation-specific genes.

The invention under consideration furthermore provides for a cell, preferably a mammalian cell, containing a gene-trapping construct such as that described above.

In a favoured operational form the mammalian cell is dependent on IL-3 and contains a gene-trapping construct in which the first reporter gene codes for Cre recombinase, whereby the Cre recombinase can delete a DNA fragment situated before the second reporter gene, where the deleted fragment codes for a thymidine kinase-neomycin phosphotransferase fusion protein and is flanked by loxP target sequences, and where the second reporter gene codes for IL-3 and following deletion of the thymidine kinase-neomycin phosphotransferase fusion protein gene comes to lie immediately downstream of the phosphoglycerate kinase promotor.

It is especially preferred that the mammalian cell is a growth factor-dependent haematopoietic precursor cell (e.g. FDCP1) or a totipotent embryonic stem cell.

According to the invention under consideration the mammalian cell can be used for identification and/or isolation of genes, especially transient genes, by which means particularly identification and/or isolation of the genes responsible for programmed cell death, of cell cycle genes, DNA repair genes and differentiation-specific genes is favoured.

Furthermore the invention under consideration provides for a vector containing the gene-trapping construct described above.

In addition a kit for the identification and/or isolation of genes, particularly transient genes, is proposed, containing at least one gene-trapping construct such as is described above.

Preferably the kit contains two constructs, by which the first construct contains the first reporter gene and the second construct the second reporter gene.

The kit may also contain a vector such as described above.

6

In conclusion, the invention under consideration provides a process for the identification and isolation of genes, especially transient genes, which contains the following steps:

- Installation of a gene-trapping construct as described above, in a suitable cell;

- Selection of cells in which the first reporter gene is incorporated in a gene, preferably an inactive gene;

- Activation of the first reporter gene, preferably by initiation of transcription of the inactive gene, in consequence of which the second reporter gene is activated;

- Identification and/or isolation of the gene in which the first reporter gene is incorporated.

- In a favoured operational form the process described above is characterized thereby, that

- a gene-trapping construct such as described above contains as first reporter gene the Cre recombinase gene, and that the Cre recombinase deletes a thymidine kinase-neomycin phosphotransferase gene situated upstream of the second reporter gene, through which the second reporter gene. which codes for IL-3, comes under the control of a phosphoglycerate kinase promotor,

- the gene-trapping construct is inserted into an IL-3-dependent FDCP-1 cell,

- through culture in neomycin- (e.g. G 418-)containing medium cells are selected for which do not produce Cre recombinase,

- these selected cells are cultivated in IL-3-free medium so that the genes responsible for programmed cell death are activated,

- that the surviving cells are selected out and

- the genes responsible for programmed cell death are isolated by the customary procedures.

- In a further favoured operational form the process is thereby characterized, that

- a gene-trapping construct, as described above, contains the Cre recombinase gene as first reporter gene, and that the Cre recombinase deletes a neomycin phosphotransferase gene situated upstream of the second reporter gene, through which the second reporter gene, which codes for lacZ, comes under the control of a phosphoglycerate kinase promotor,

- the gene-trapping construct is inserted into a totipotent embryonic stem cell,

- through cultivation on G418 cells are selected for which do not produce Cre recombinase,

- these selective cells are induced to differentiate, so that differentiation-specific genes are activated,

- these cells are isolated,

- these cells are introduced into the germ line of a mammal and are responsible there for the shaping of tissues, and

- the differentiation-specific genes are isolated by the customary procedures.

The abovementioned process for the isolation of the genes responsible for programmed cell death is for this reason especially appropriate, since through permanent replacement of selection markers by a gene trap expressing Cre recombinase, they are converted through activation in essentially dying cells of the endogenous cytokine production into multiplying cells. This permits selection for integration into genes which are active during the programmed cell death. Expression of the selection marker, for instance IL-3, is disassociated from the expression of the localized gene (in the gene trap) by which the recombinase brings the selection marker behind the phosphoglycerate kinase promotor. This is especially important as it is improbable that cells should go on expressing the genes responsible for programmed cell death. Finally, the process according to the invention under discussion is not confined to strongly expressed genes, and the gene representation is very much more uniform than for example in cDNA cloning. Furthermore the use of gene traps for isolation and/or identification, unlike differential RT-PCR amplification, is well reproducible, avoids redundancy and even enables quantification. In conclusion, the gene trap proviruses are regularly

located in or near the 5' exon and have the same transcriptional orientation relating to the gene, and this clearly facilitates cloning.

In any case characterization of the cellular transcripts induced by IL-3 withdrawal makes possible deeper insights into the complex molecular changes taking place during programmed cell death.

In a further favoured operational form a totipotent cell, which in itself can develop into an entire organism, contains a construct whereby the first reporter gene codes for a Cre recombinase which can delete a fragment situated before the second reporter gene, whereby the deleted fragment codes for neo and is flanked by loxP target sequences, and the second reporter gene codes for lacZ which following deletion of neo comes to lie immediately downstream of pgk.

The appended figures should explain the invention under discussion:

Figure 1 shows the process relating to the invention in a schematic representation of two favoured operational forms.

(A) Activation mechanism of the endogenous IL-3 secretion in FLOXIL-3 cells after IL-3 withdrawal. Integration of a U3-Cre gene trap into an originally unexpressed apoptosis gene is transiently activated by IL-3 withdrawal (left). Through site-specific recombination of the reporter plasmid ppgklxtkneoIL3 the tkneo gene is eliminated and the IL-3 gene is placed immediately downstream of the pgk promotor. This results in sustained IL-3 synthesis which continues even after switching off of the gene captured in the gene trap.

(B) Activation mechanism of the lacZ gene during differentiation of -pln13- totipotent embryonic stem cells. A U3-Cre gene trap integration into an originally unexpressed differentiation gene is transiently activated during the differentiation. Through site-specific recombination of the reporter plasmid ppgklxneoLacZ the LacZ gene is placed immediately downstream of the pgk promotor. This results in sustained synthesis of beta-galactosidase, which continues even after switching off of the gene captured in the gene trap.

Figure 2 shows a schematic overview of a process for the isolation and analysing of inducible differentiation genes from embryonic stem cells.

Figure 3 shows the Cre/loxP-mediated recombination of the ppgklxtkneoIL3 plasmid in FLOXIL-3 cells with U3-Cre-gene trap integrations in expressed genes.

(A) Structure of ppgklxtkneoIL3 before and after recombination.

(B) Southern-Blot analysis of autonomous and parental FLOXIL-3 cells. Genomic DNAs were cut with BamHI, fractionated in agarose gels and transferred on to nylon membranes. Hybridization followed with ^{32}P-labelled neo-(left) or pgk-specific probes. Lane 1: parental FLOXIL-3 cells; Lanes 2-6: autonomous clones 1-5. Molecular weights are expressed with the help of a 1-kb BRL scale.

Figure 4 shows U3-Cre gene trap integrations and site-specific recombination in autonomous clones isolated after withdrawal of IL-3.
(Above) Expected structure of U3-Cre proviruses (left) and the recombined reporter plasmid (right).
(Below) Southern-Blot analysis of clones isolated following IL-3 withdrawal from the U3-Cre/Floxil-3 integration bank. The genomic DNA of individual clones was cut with BamHI and treated as in the text of Figure 3. Hybridization followed with ^{32}P-labelled Cre- (left) or pgk- (right) specific probes. Lane 1, 26-11-1; lane 2, 26-11-3; lane 3, 26-11-4; lane 4, 26-11-5; lane 5, 26-11-6; lane6, 26-11-7; lane 7, 26-11-8; lane 8, 26-12-1; lane 9, 26-12-2; lane 10, 26-12-3; lane 11, 26-12-4; lane 12, 26-12-5. The probes illustrated are representative of the whole integration bank.

Figure 5 shows an investigation of the cell-provirus fusion transcript.
(Above) Transcript expected from U3-Cre gene-trap integration in expressed genes.
(Below) Northern-Blot analysis of cell-provirus fusion transcripts before and after serum withdrawal. Polyadenylated RNA (5 micrograms) was fractionated in formaldehyde-agarose gels and transferred on to nylon membranes. Hybridization followed with ^{32}P-labelled Cre- or GAPDH-specific probes. Odd numbers stand for RNAs, even numbers RNAs after 16-hour serum withdrawal. Lanes 5 and 6 contain RNA from a clone with a constitutively expressed U3-Cre integration.

Figure 6 shows an investigation of cellular transcripts following IL-3 withdrawal. Polyadenylated RNA (5 micrograms per lane) was derived from FDCP-1 cells deprived of IL-3 for 0, 6 and 12 hours. Northern-Blots were constructed as in the text of Figure 5 and hybridized with ^{32}P-labelled 5' provirus-flanking sequences or GAPDH.

Figure 7 shows the differentiation-regulated U3-Cre gene trap integrations into embryonic stem cells. 960 pools each of 150 cells were infected with the U3-Cre virus and selected in G418. Differentiation induction followed over 4 days.
(Left) X-Gal staining before (left) and after (right) induction of differentiation.

(Right) Southern-Blot analysis of the regulated clones. The genomic DNA was cut with BamHI and treated as in the text of Figure 3. Hybridization followed with ^{32}P-labelled Cre- (left) and pgk- (right) specific probes.

u = undifferentiated, d = differentiated.

Particulars of this are to be found in the Examples.

In the following the invention will be described in detail with the help of examples; the examples will not however be limited to the scope of the invention.

EXAMPLES

Example 1

Construction of a reporter plasmid

Most of the components of the reporter plasmid ppgklxtkneoIL-3 (phosphoglycerate transferase promotor, loxP target sequence, thymidine kinase-neomycin phosphotransferase fusion protein, IL-3 gene) were put together in p-BluescriptII-KS as follows:

The sequences for the target sequence loxP were derived from pGEM30. A loxP site was inserted at first into the Bluescript Polylinker as an EcoRI/Psti fragment. Then pgk (phosphoglycerate kinase promotor, obtained from ppgkCAT) was ligated to the XbaI/BamHI recombination site of the polylinker and the thymidine kinase-neomycin phosphotransferase fusion gene was cloned in the EcoRV recombination site situated downstream. The IL-3 gene (as cDNA of mice, from pc-Multi-CSF) was first subcloned from pGEM30 in the EcoRI site of the polylinker flanking the loxP target sequence, then next the plasmid was cut with ClaI/SalI and the endfilled ends were again ligated to remove an additional EcoRI site. IL-3-loxP was then recovered from pGEM30 as an SalI/XhoI fragment and cloned in the Bluescript Polylinker. A copy of the bovine growth hormone polyadenylation sequence (bpa) as XbaI/AvaI fragment was cloned with blunt ends in the ClaI site downstream of tkneo. In order to obtain ppgklxtkneoIL-3 a XhoI fragment containing the assembled sequences was cloned in pSBC2 upstream of an SV-40 polyadenylation site. Since this original construct still enabled translation of IL-3 from the dicistronic tkneo-IL-3 transcript, a second copy of

bpa with blunt ends was cloned in the SalI site of pSBCII in order to obtain the definitive ppgklxtkneolL-3 reporter plasmid.

Example 2

Construction of an FDCP1 reporter cell line expressing two selectable reporter genes flanked by loxP recombination targets

The FDCP1 haemopoietic precursor cells used are strictly dependent on IL-3 for their growth and initiate programmed cell death if this growth factor is withdrawn from the medium.

The FDCP1 cells were cultivated at concentrations of 2×10^5 cells per ml in Dulbeccos modified Eagle's medium (DMEM: Gibco) supplemented with 10% v/v foetal calf serum (High Clone Laboratories, Utah, USA) and 10ng/ml recombinant mouse IL-3 (Sandoz). The agar cultures were a mixture of equal volumes of double-strength DMEM supplemented with 40% (v/v) foetal calf serum (Hyclone) and 0.6% (w/v) Bactoagar (Difco) in twice-distilled water. A few cultures contained 5 muM ganciclovir (Syndex).

In order to obtain a cell line which is transformed by Cre recombinase into a factor-independent cell, FDCP1 cells were electroporated with the reporter plasmid ppgklxtkneolL-3. The electroporation was carried out using a BioRad Gene Pulser* (BioRad) following the instructions of the manufacturer. The recombinants were isolated in agar culture containing 0.6 mg/ml G418. After 10 days the developing clones were isolated and amplified in suspension cultures as described above. The ppgklxtkneolL-3 (Figure 3A) consists of two selectable reporter genes arranged one after the other, which are transcribt from a pgk (phosphoglycerate kinase) promotor. The gene at the 5' end codes for an HSV2 thymidine kinase (Tk) neomycin phosphotransferase (neo) fusion protein and is flanked by two loxP recombination targets. The 3' end of the gene codes for mouse IL-3 and terminates in an SV40 polyadenylation sequence. To suppress the IL-3 translation of dicistronic transcripts, two consecutively arranged copies of the polyadenylation sequence of bovine growth hormone (bpa) were inserted downstream from tkneo. In this way one could assume that even cells which expressed ppgklxtkneolL-3 were still always IL-3 dependent. Since the Cre recombinase typically excises those sequences that are flanked by direct repetitions of lox-P, it was assumed that expression of Cre deleted the tkneo gene and thereby placed the IL-3 gene

*Trade-mark

downstream of the pgk promotor. As a result the cells would lose their neomycin resistance and acquire factor independence through synthesis of IL-3.

Stable FDCP-1 transformants were selected in G418 and five clonal cell lines were isolated from the agar cultures. A cell line expressing a single-copy plasmid and in addition identified as FLOXIL-3 was selected for further analysis.

It was first established whether the FLOXIL-3 cells were still dependent for their growth on IL-3, by plating out 5×10^7 cells at a concentration of 2×10^5 per ml in half-set agar cultures without IL-3. Since no colonies developed within 10 days, it was assumed that neither limited IL-3 translation nor spontaneous recombination had occurred inside these cells.

Further to confirm this, FLOXIL-3 and parental FDCP-1 cells were pre-incubated in agar cultures without IL-3 for up to 24 hours and finally rescued by giving IL-3. Both cell types initiated programmed cell death with similar kinetics, which indicated that the expression of ppgklxtkneoIL-3 did not alter the cellular response to the factor withdrawal. Moreover most of the cells survived for more than 12 hours without IL-3 and so left enough time for the expression of the sequences transduced by the gene trap.

Example 3

<u>Construction of a gene trap which expresses Cre recombinase (U3-Cre) and renders FLOXIL-3 cells factor-independent</u>

A U3-Cre gene-trap vector was derived from the pGgU3Neo(en-) vector based on MoMuLV, by means of which the neo gene in the U3 region was replaced with the sequence coding for Cre derived from pMCCre. The plasmid was transfected into helper cells, and the remainder of the cell line, which produced recombinant viruses in high titre, was used for infecting FLOXIL-3 cells. As already indicated earlier, the virus replication and LTR-mediated duplication of sequences inserted into U3 places precisely 30 nucleotides downstream of the flanking cellular DNA. This enables their expression through integrations into transcribed genes. It was expected from this that FLOXIL-3 cells with U3-Cre integrations in expressed genes would be transformed into factor-independent cells. In order to select for this occurrence, virus-infected FLOXIL-3 cells were plated out in agar cultures without IL-3. Many autonomously growing colonies were obtained and amplified in

suspension cultures. As was to be expected from the expression of a recombinant reporter plasmid, all the clones lost their G418 resistance and grew without IL-3.

In order to verify the recombination, genome DNAs from five autonomous clones were digested with BamHI and analysed by Southern blotting. Since BamHI divides inside the 5' ends of pgk and IL-3, recombinant FLOXIL-3 cells do not replicate an internal fragment of 3.2 kb which hybridizes both with pgk- and with neo-specific probes (Figure 3A). Since this fragment contains all the sequences which are flanked by loxP, including tkneo, a deletion should be associated with a reduction in size of 2.6 kb. In fact all clones produced a 0.6 kb restriction fragment which did not hybridize with neo, which indicated that recombination had taken place (Figure 3B).

Furthermore, when analysed by Northern blotting all the clones expressed cell provirus transcripts, which indicated that the Cre recombinase of active cellular promotors was being expressed.

In this way FLOXIL-3 cells are obtained containing a gene-trapping construct which enables such genes as are reponsible for programmed cell death to be isolated and identified.

Example 4

Construction of a U3-Cre/FLOXIL-3 integration bank and isolation of transiently expressed gene sequences

An integration bank consisting of 2 x 10^6 independent proviral integrations was constructed by infection of FLOXIL-3 cells with U3-Cre retroviruses. The infected cells were first preselected in G418 in order to eliminate complete integrations into expressed genes. Figure 4 shows that this goal was attained following 16 days' selection. Following plating out of a basic sample of the integration bank in agar culture without IL-3, altogether 110 autonomous clones could be isolated and amplified. Southern blot investigations have shown that all the clones displayed provirus integration and without exception contained recombinant reporter plasmids (Figure 4).

Northern blot investigations with mRNA from 11 independent autonomous clones show that in 9 clones cell-provirus fusion transcripts were either very weakly expressed or not at all (Figure 5, lanes 1, 3, 7, 9, 11, 13, 17, 19, 21). This means that the Cre recombinase was only transiently expressed sufficiently to cause recombination of the reporter plasmid. In addition more than 50% of

these fusion transcripts were inducible by serum withdrawal, which indicates, that the genes
intercepted by the gene trap are associated with arrest of an apoptotic programme or of growth
(Figure 5).

Finally the 5'-provirus-flanking sequences cloned by known processes (von Melchner et al., 1992)
were hybridized with the mRNA of wild-type FDCP-1 cells on Northern blots. The majority of the
probes investigated hybridized with IL-3-induced cellular transcripts (Figure 6). Whether these
genes control the apoptosis directly or are part of other mechanisms which are activated as a result
of the programmed cell death, still has to be established. In any case characterization of the cellular
transcripts induced by the IL-3 withdrawal facilitates further insights into the complex molecular
changes which emerge during programmed cell death.

The further precise identification and isolation of the discovered genes can be carried out by
customary processes, such as for instance 3'RACE or the restoration of a gene library and cloning
and sequencing of the genes discovered.

Example 5

Construction of an ES reporter cell line

In order to obtain an ES cell line suitable for a permanent selectable reporter exchange through Cre
recombinase, D3 cells were electroporated with the reporter construct ppgklxneolacZ in which two
reporter genes arranged one after the other were transcribed by a mouse phosphoglycerate kinase
promotor. The reporter construct ppgklxneolacZ consists of the phosphoglycerate kinase promotor,
the loxP target sequences, the neomycin phosphotransferase gene and a gene for lacZ. Its
construction followed the pattern described above for construction of the reporter plasmid
ppgklxtkneolL-3.

The 5' gene codes for neomycin phosphotransferase and is flanked by two loxP recombination
targets in identical orientation. The 3' gene codes for beta-galactosidase and ends in an SV40
polyadenylation sequence. In order to suppress genomic transcripts ending in SV40-PolyA and in
this way also to suppress the translation of lacZ, two consecutively positioned copies of the bovine
growth hormone polyadenylation sequence (bpa) are inserted downstream of the neomycin
phosphotransferase. Cells expressing ppgklxneolacZ should accordingly display a phenotype

resistant to neomycin phosphotransferase and negative for lacZ. In the case of Cre expression, however, the neo gene would be excised from the recombinase and the sequence coding for lacZ placed immediately downstream of the pgk promotor. As a result the cells which express the recombined reporter plasmid lose their neomycin resistance and stain positive with X-Gal (lacZ+).

Various ES cell lines expressing ppgklxneolacZ were isolated in G418 and tested individually on a background of X-Gal staining. Four of 10 cell lines did not stain regularly with X-Gal even after long periods of culture. If one induced these cell lines to differentiate in cultures containing no LIF (Leukaemia Inhibiting Factor) or nutrient layers, lacZ cells still stained negative, which indicated that the lacZ phenotype is stable. A cell line expressing a single copy of ppgklxneolacZ was selected for further analysis. In the following it is named pln13.

In order to establish whether the Cre recombinase converted the lacZ of pln13 into lacZ+, the cells were transfected with the Cre-expressing plasmid pMC-Cre. Unselected colonies were picked after 10 days and aliquots stained with X-Gal. Genomic DNA extracted from various lacZ+ cells was cut with BamHI and analysed by Southern blotting. In every case the lacZ phenotype was associated with a 2 kb deletion, which indicated that the reporter plasmid was subject to site-specific recombination.

Example 6

Expression of beta-galactosidase in pln13 cells through U3-Cre integration

5×10^4 pln13 cells were infected at MOI = 1 with U3-Cre virus constructed as described above. After incubation for 72 hours the cells were divided into 480 pools and seeded out in microtitre plates. After 48 hours aliquots were stained with X-Gal in order to detect integrations into active cells. Out of 20 pools containing lacZ+ cells three were seeded out at clonal density. Four positive clones were selectively isolated and analysed by Southern blotting. All clones contained 1-2 proviruses which agreed with an MOI = 1 and contained recombined reporter plasmids. Every clone in addition expressed cell-provirus fusion transcripts whose size in accordance with expectations varied around 100 to 500 nt.

In this way U3-Cre gene trap retroviruses effectively mediate site-specific recombinations on account of their integration into expressed genes.

Example 7

Isolation of clones with inducible U3-Cre-provirus integration

In order to establish whether the U3-Cre/loxP strategy is suitable for isolating genes which owing to differentiation are inducible, 1.5×10^5 pln13 cells were seeded out in the cups of 10 microtitre plates at concentrations of 150 cells per cup and infected with U3-Cre virus at about MOI = 1. In order to eliminate integration into active genes, cells were selected on G418 for 5 days. Then in the absence of either LIF (Leukaemia Inhibitory Factor) or MEF (Mouse Embryonic Fibroblasts) aliquots of each cup were induced to differentiate and stained with X-Gal. Forty-four out of 960 cups stained positive by this test. Following a further test 9 pools were lacZ+ in both presence and absence of LIF and MEF, while 35 pools changed to lacZ+ only following differentiation, which indicated that U3-Cre integration had occurred in the induced genes.

Figure 7 shows as an example three isolated clones with U3-Cre-provirus integrations into differentiation-specific genes. The inducibility of the genes was attested by the phenotypic change from LacZ(-) (Figure 7A, undifferentiated cells, above left) to LacZ(+) (Figure 7A, differentiated cells, above right). The Southern Blots indicate that this change is based on site-specific recombination of the reporter plasmid, caused by the U3-Cre-provirus integration into inducible differentiation genes (Figure 7B).

In Figure 2 is shown an overview of an isolation of this kind. In the first step come infection with U3-Cre gene trap vector and selection on G418. The colonies obtained are divided into aliquots, induced to differentiate and stained with X-Gal. Only the stained clones are selected further, amplified and used in the construction of chimaeric mice. These mice may on the other hand be used to determine the LacZ distribution in the tissue, or to determine the phenotype of F2 mice and to isolate -/-ES- cells.

The corresponding cellular genes can be cloned from the known gene-trap sequences by conventional means.

In this manner the gene trapping construct according to the invention under consideration makes it possible to isolate and/or identify genes, in particular genes only transiently expressed.

17

Bibliography

Coffin et al. (1989), "Genetics of Endogenous Murine Leukemia Viruses," Ann. NY Acad. Sci., vol. 567:39-49.

Goff (Jun. 1990), "Integration of Retroviral DNA into the Genome of the Infected Cell," Cancer Cells, vol. 2(6):172-178.

Gridley et al. (Jun. 1987), "Insertional Mutagenesis in Mice," TIG, vol. 3(6):162-166.

Jaenisch et al. (1985), "Retroviruses and Insertional Mutagenesis," Cold Spring Harbor Symp. Quant. Biol., vol. 50:439-445.

Hill et al. (1993), "Gene and Enhancer Trapping: Mutagenic Strategies for Developmental Studies," Curr. Top. in Dev. Biol., vol. 28:181-206.

Hill et al. (1993), "Screening for Novel Pattern Formation Genes Using Gene Trap Approaches," Methods in Enzymology, vol. 225:664-681.

Muthuchamy et al. (Jun. 1993), "Developmental Analysis of Tropomyosin Gene Expression in Embryonic Stem Cells and Mouse Embryos," Molecular and Cellular Biology, vol. 13(6):3311-3323.

Rappolee et al. (Sep. 1988), "Developmental Expression of PDGF, TGF-.alpha., and TGF-.beta. Genes in Preimplantation Mouse Embryos," Science, vol. 241:1823-1825.

Reddy et al. (Aug. 1992), "Fluorescence-Activated Sorting of Totipotent Embryonic Stem Cells Expressing Developmentally Regulated lacZ Fusion Genes," Proc. Natl. Acad. Sci. USA, vol. 89:6721-6725.

Rogers et al. (1991), "Specific Expression of a Retinoic Acid-Regulated, Zinc-Finger Gene, Rex-1, in Preimplantation Embryos, Trophoblast and Spermatocytes," Development, vol. 113:815-824.

Sandmeyer et al. (1990), "Integration Specificity of Retrotransposons and Retroviruses," Ann. Rev. Genet., vol. 24:491-518.

Scholer et al. (1990), "Oct-4: A Germline-Specific Transcripton Factor Mapping to the Mouse t-Complex," EMBO Journal, vol. 9(7):2185-2195.

Sharpe et al. (Dec. 1990), "Regulated Expression of the Small Nuclear Ribonucleoprotein Particle Protein SmN in Embryonic Stem Cell Differentiation," Molecular and Cellular Biology, vol. 10(12):6817-6820.

CLAIMS

1.   Gene-trapping construct, containing a first reporter gene, which after activation can activate a second reporter gene,

wherein the first reporter gene codes for a recombinase;

wherein the second reporter gene is activated thereby, that the recombinase · deletes a DNA fragment located before the second reporter gene and in that way places the second reporter gene downstream from a promotor, said second reporter gene being under control of said promotor; and

wherein the deleted DNA fragment is flanked by target sequences for the recombinase.

2.   Gene-trapping construct according to claim 1, characterized thereby that the recombinase is Cre.

3.   Gene-trapping construct according to claim 1, characterized thereby that the recombinase is Flp.

4.   Gene-trapping construct according to any one of claims 1 to 3, characterized thereby that the second reporter gene codes for a protein factor.

5.   Gene-trapping construct according to claim 4, characterized thereby that the second reporter gene codes for IL-3.

6.   Gene-trapping construct according to claim 4, characterized thereby that the second reporter gene codes for LacZ.

7.   Gene-trapping construct according to any one of claims 1 to 6, characterized thereby that the deleted fragment is an antibiotic-resistance gene.

8.   Gene-trapping construct according to claim 7, characterized thereby that the deleted DNA fragment codes for a thymidine kinase-neomycin phosphotransferase fusion protein.

9.   Gene-trapping construct according to claim 8, characterized thereby that the deleted DNA fragment codes for a neomycin phosphotransferase.

10.   Gene-trapping construct according to any one of claims 7 to 9, characterized thereby that the promotor is the phosphoglycerate kinase (pgk) promotor.

11.   Gene-trapping construct according to any one of claims 1 to 10, characterized thereby that the target sequence is loxP.

12.   Gene-trapping construct according to any one of claims 1 to 10, characterized thereby that the target sequence is frt.

13.   Gene-trapping construct according to any one of claims 1 to 12, characterized thereby that the target sequence is located in a U3 and/or a U5 region.

14.   Use of the gene-trapping construct according to any one of claims 1 to 13 for identification and/or isolation of genes.

15.   Use according to claim 14 for identification and/or isolation of genes controlling programmed cell death, DNA repair, differentiation and/or the cell cycle.

16.   Cell containing a gene-trapping construct according to any one of claims 1 to 13.

17.   Cell according to claim 16 characterized thereby that said cell is mammalian and it is dependent on IL-3, that the first reporter gene codes for Cre recombinase, that the Cre recombinase can delete a DNA fragment located before the second reporter gene, that the deleted fragment codes for a thymidine kinase-neomycin transferase fusion protein (TKNeoPT) and is flanked by loxP target sequences, that the second reporter

gene codes for IL-3 and following deletion of the TKNeoPT gene comes to lie downstream of a phosphoglycerate kinase promotor.

18.   Cell according to claim 17, characterized thereby that it is a haemopoietic FDCP-1 cell or a totipotent stem cell.

19.   Use of the cell according to claim 16 for identification and/or isolation of genes, wherein said cell is a mammalian cell.

20.   Use of the cell according to any one of claims 16 to 18 for identification and/or isolation of genes controlling programmed cell death, DNA repair, differentiation and/or the cell cycle.

21.   Vector containing the gene-trapping construct according to one or more of. claims 1 to 13.

22.   Process for the identification and/or isolation of genes comprising the following steps:

        introduction of a gene-trapping construct according to one of the claims 1 to 13 into a suitable cell;

        selection of cells in which the first reporter gene is incorporated in a gene;

        activation of the first reporter gene through initiation of the transcription of the first reporter gene, in consequence of which the second reporter gene is activated; and

        identification and/or isolation of the gene in which the first reporter gene is incorporated.

23.   Process according to claim 22, characterized thereby that

        a gene-trapping construct according to claim 1 contains as first reporter gene the Cre recombinase gene, that the Cre recombinase deletes a thymidine kinase-neomycin phosphotransferase gene situated downstream from the second reporter gene, whereby the second reporter gene, which codes for IL-3, comes under the control of a phosphoglycerate kinase promotor;

the gene-trapping construct is inserted into an IL-3-dependent FDCP-1 cell;

through culture in a neomycin medium cells which do not produce Cre recombinase are selected;

these selected cells are cultivated in IL-3-free medium, so that the genes responsible for the programmed cell death are activated;

that the surviving cells are selected out; and

genes responsible for programmed cell death are isolated by the customary procedures.

24.   Process according to claim 22, characterized thereby that:

a gene-trapping construct according to claim 1 contains the Cre recombinase gene as first reporter gene, that the Cre recombinase deletes a neomycin phosphotransferase gene located downstream from the second reporter gene, whereby the second reporter gene, which codes for lacZ, comes under the control of a phosphoglycerate kinase promotor;

the gene-trapping construct is inserted into a totipotent embryonic stem cell;

through culture in G418, cells which do not produce Cre recombinase are selected;

these selected cells are induced to differentiate, so that differentiation-specific genes are activated;'

these cells are isolated;

these cells are introduced into the germ line of a mammal and are there responsible for the formation of tissues; and

the differentiation-specific genes are isolated by the customary procedures.

25.   Use according to claim 14 or 15, wherein said genes are transient genes.

26.   Use according to claim 14, wherein said genes are transient genes.

27.   Process according to any one of claims 22 to 24, wherein said identification and/or isolation of genes is the identification and/or isolation of transient genes.

28.   Process according to any one of claims 22 to 24 and 27, wherein said first reporter gene is incorporated in an inactive gene.

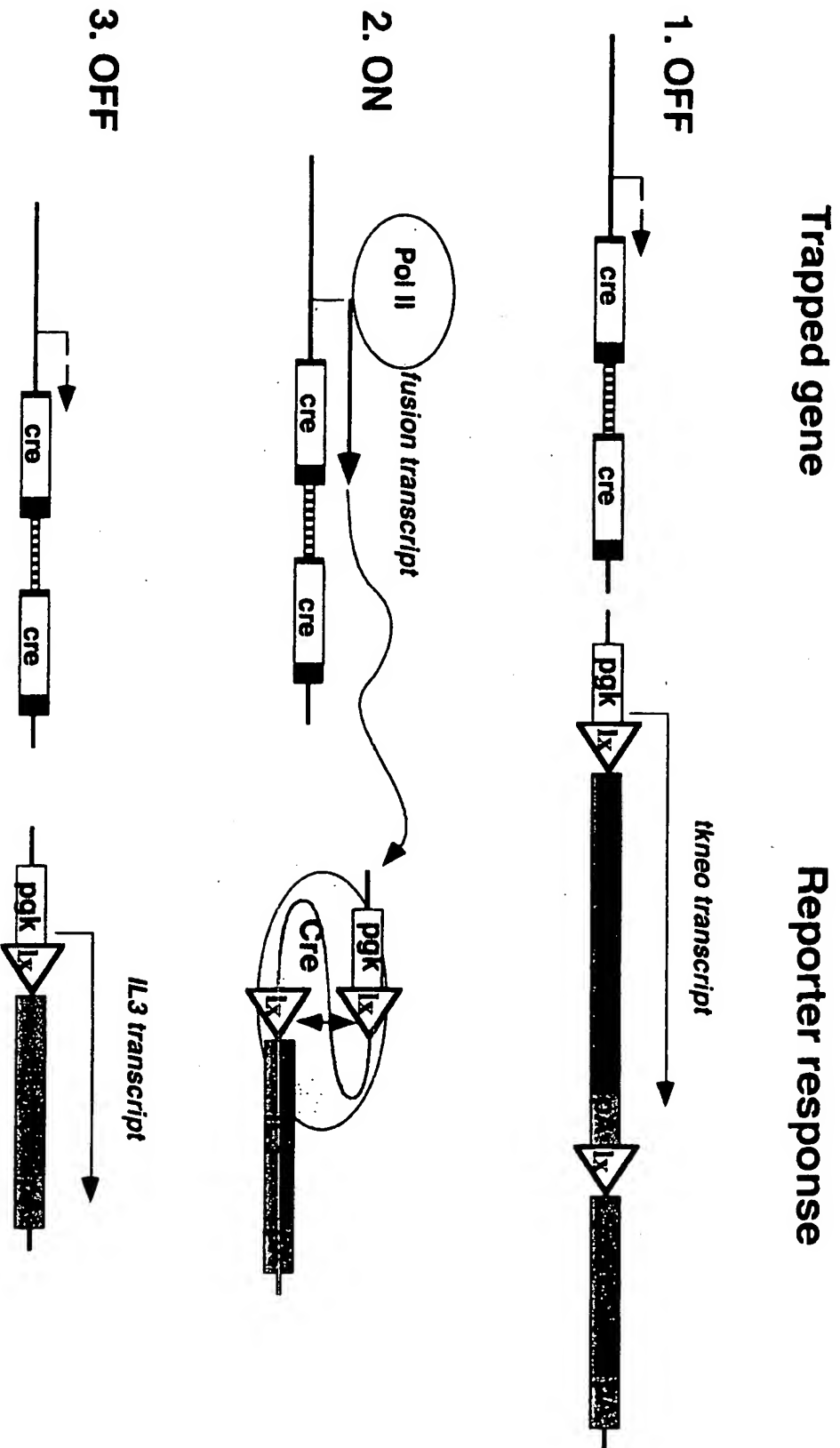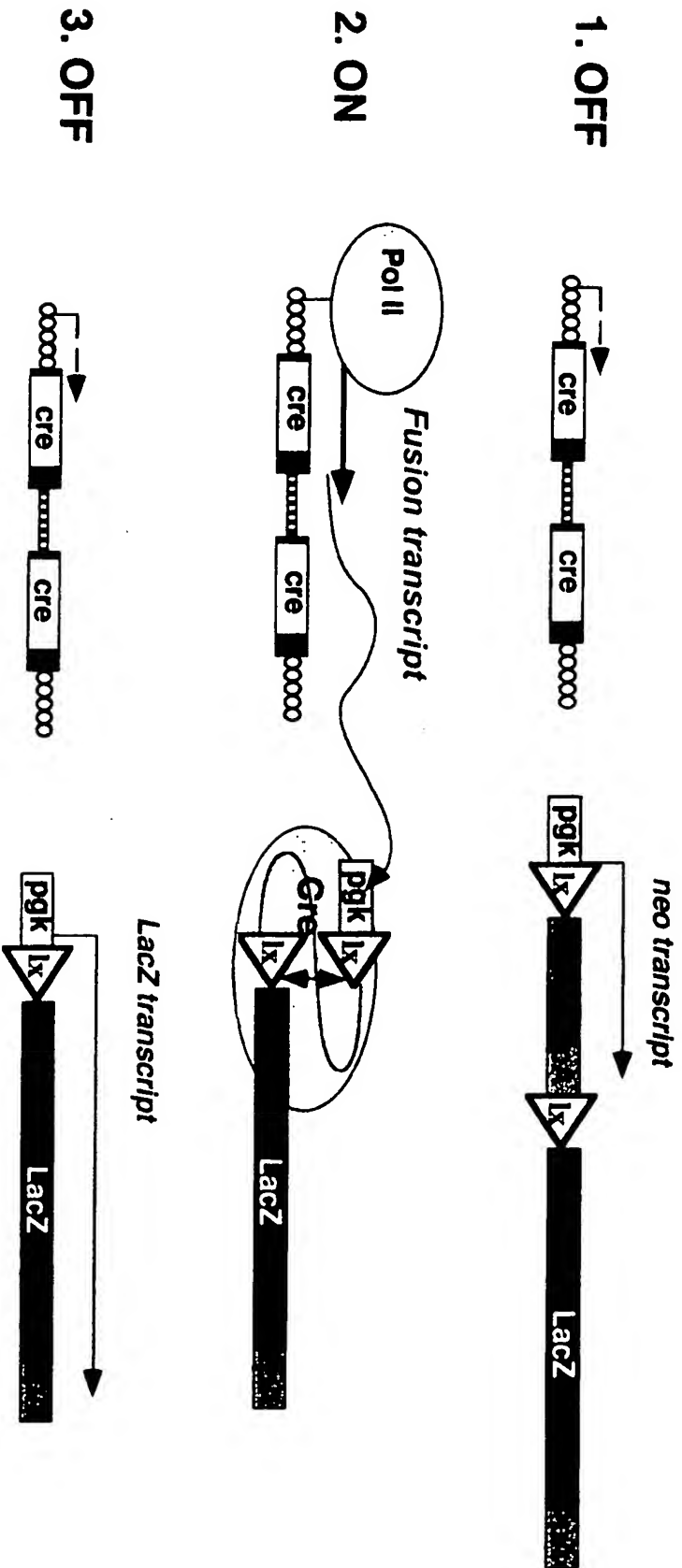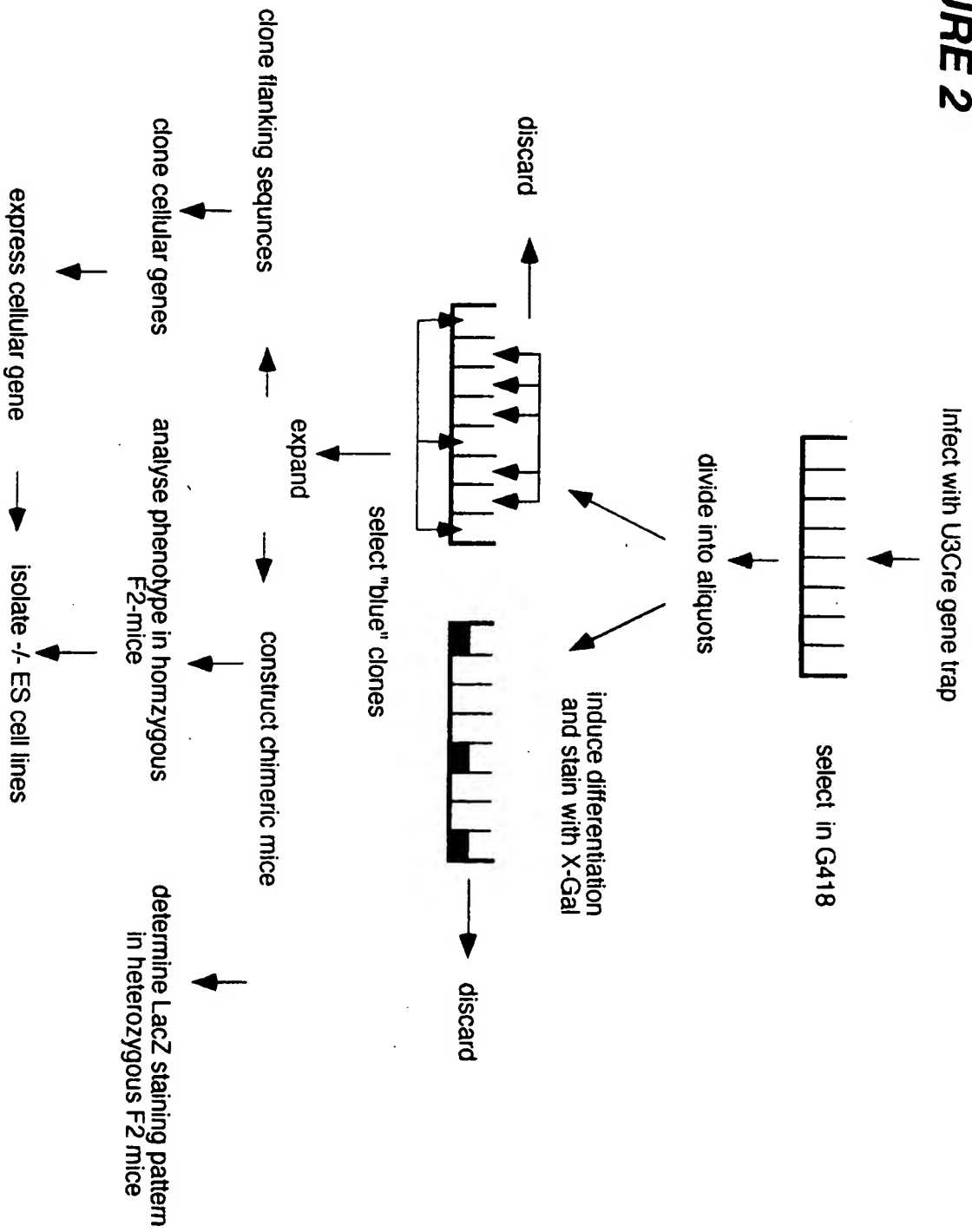29.   Cell according to claim 16, wherein said cell is a mammalian cell.
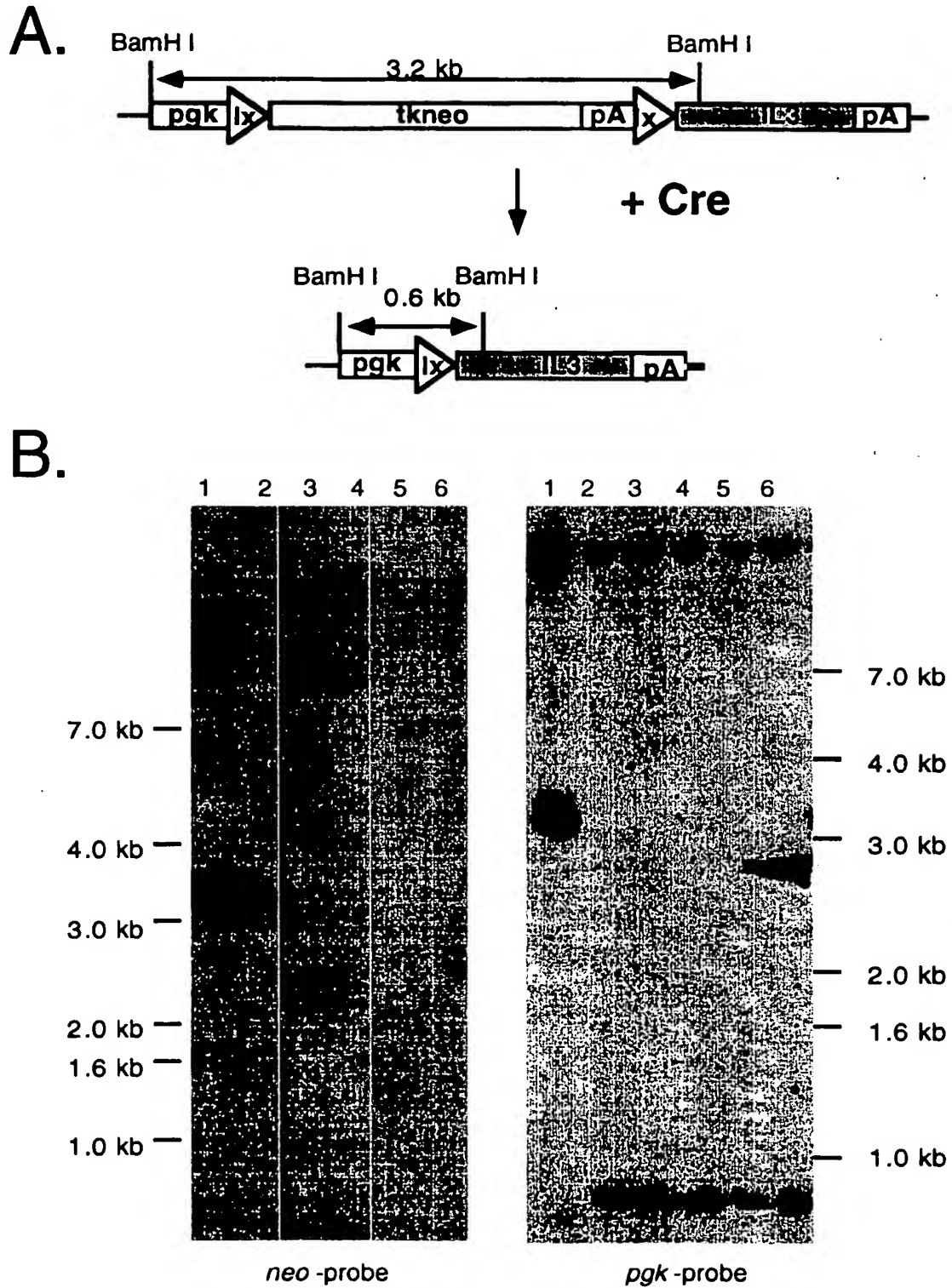
FIGURE 1A

FIGURE 1B

Trapped gene
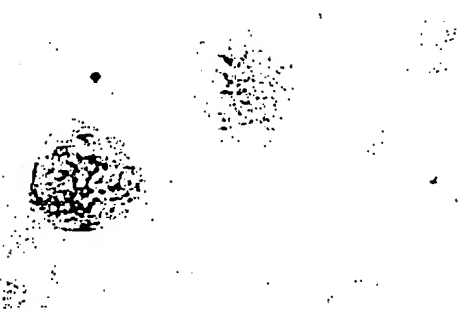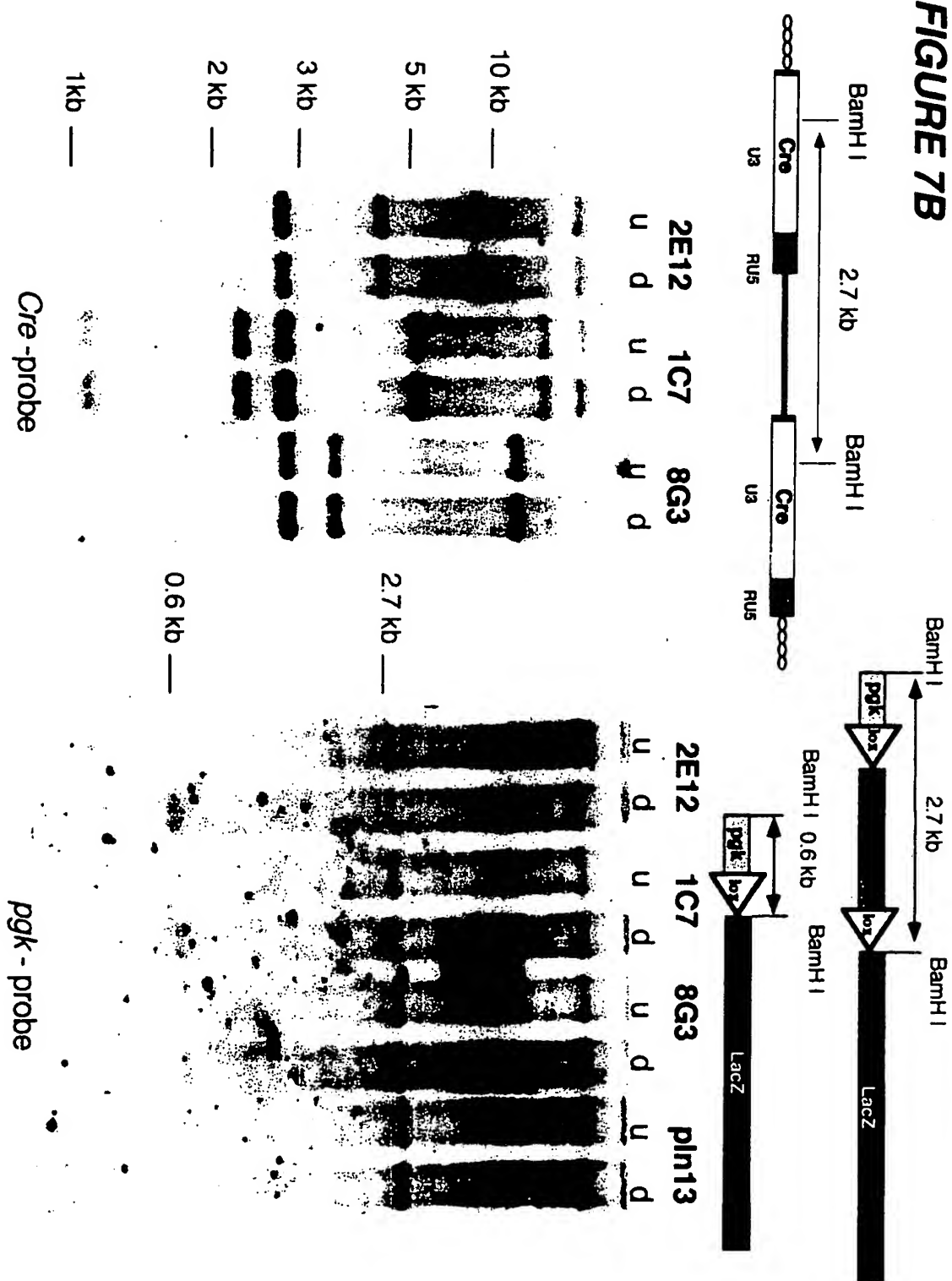
Reporter response

1. OFF

2. ON

3. OFF

FIGURE 2

Infect with U3Cre gene trap

select in G418

divide into aliquots

induce differentiation
and stain with X-Gal

discard

discard

select "blue" clones

expand

analyse phenotype in homzygous
F2-mice

construct chimeric mice

isolate -/- ES cell lines

determine LacZ staining pattern
in heterozygous F2 mice

clone flanking sequnces

clone cellular genes

express cellular gene

FIGURE 3

A.

B.

neo -probe   pgk -probe

FIGURE 4

Cre- probe

pgk- probe
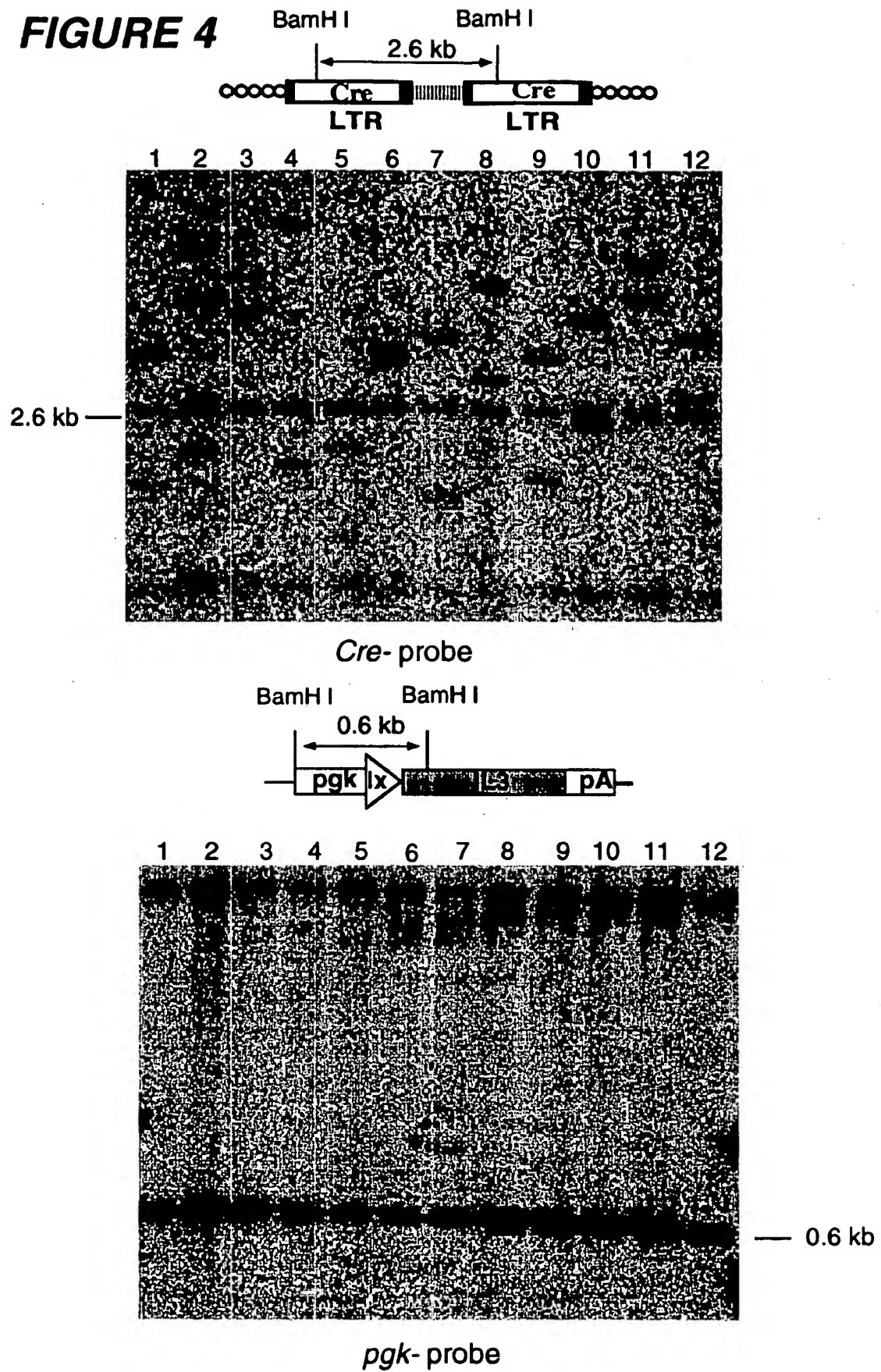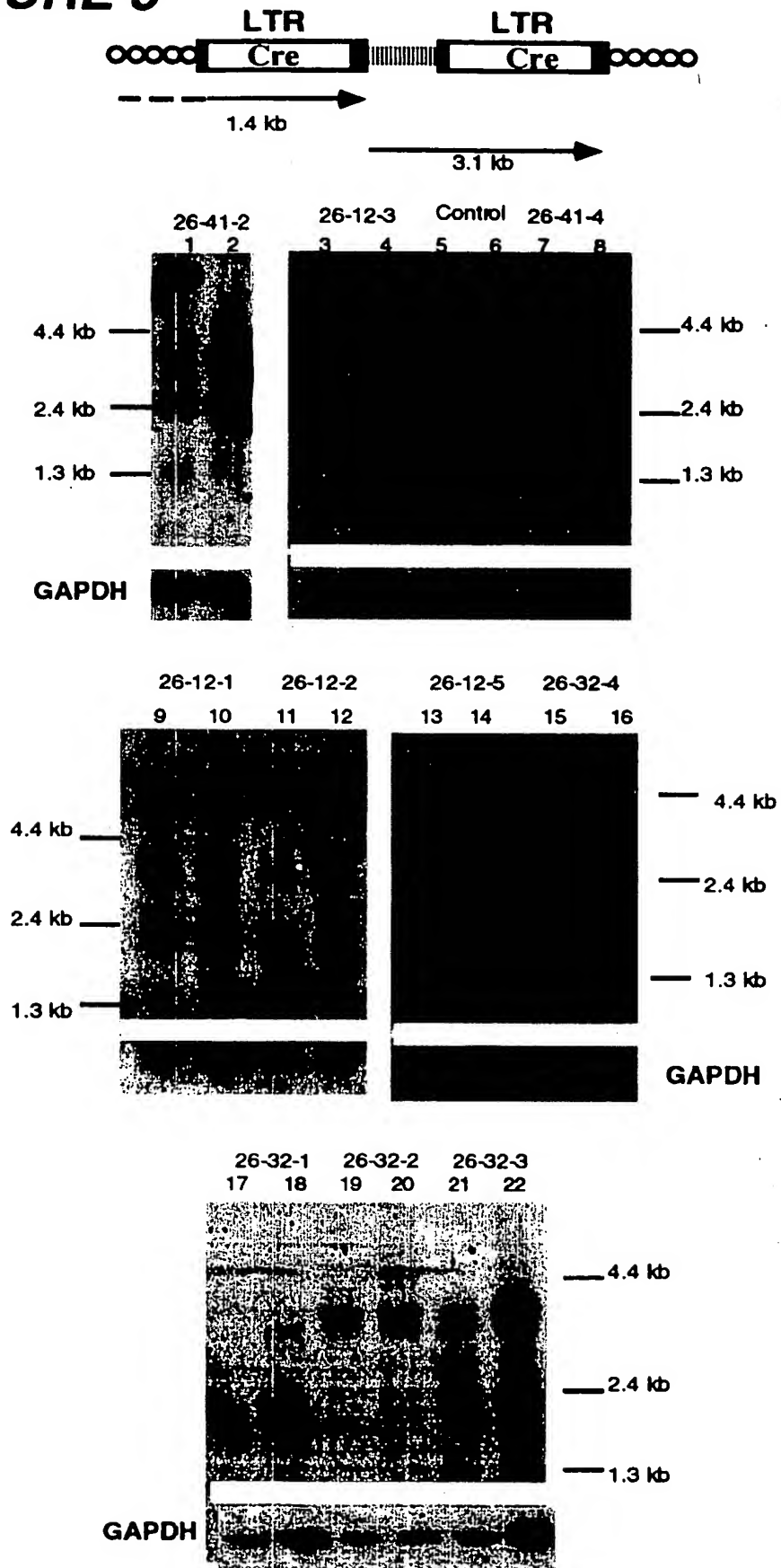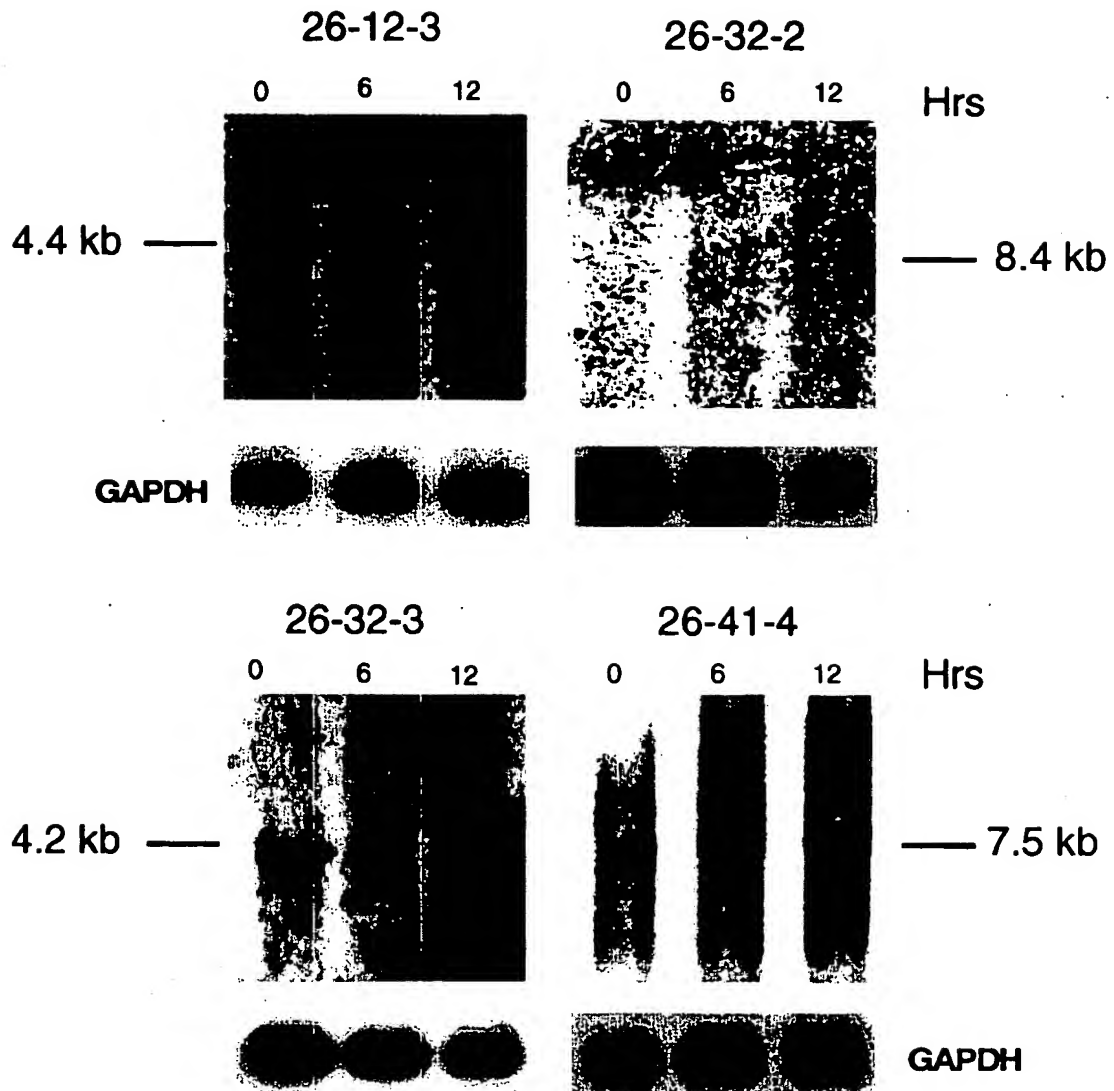
FIGURE 5

FIGURE 6

FIGURE 7A

2E12

1C7

8G3

FIGURE 7B

Trapped gene   **Reporter response**

1. OFF

tkneo transcript

pgk | lx | cre | cre | lx

2. ON

Pol II   *fusion transcript*

cre | cre

pgk | lx

Cre

lx

3. OFF

IL3 transcript

cre | cre

pgk | lx